The Company You Keep: Using Organizational Social Network Attributes to Predict Employee Engagement

1st Author Name
Affiliation
City, Country
e-mail address

2nd Author Name
Affiliation
City, Country
e-mail address

3rd Author Name
Affiliation
City, Country
e-mail address

ABSTRACT
Employee engagement (EE) has been shown to have important implications [1] for the success of organizations. Most researchers have discussed employee engagement in terms of factors in a top-down, hierarchical model of the organization. However, there may also be contributing factors from an employee’s organizational social network (OSN). In this paper, we show that an employee’s social network attributes can contribute to the prediction of engagement, primarily through centrality and homophily in a large, multinational company (LargeCorp). Our research expands the range of theoretical factors that can predict employee engagement from a vertical, three factor model to a more horizontal, four factor model. We also discuss how this work points toward a richer set of methods to predict and track engagement over time and its connections to the broader HCI research community.

Author Keywords
Employee engagement; Social networks; Homophily;

ACM Classification Keywords
H.5.3. Group and organization interfaces;

INTRODUCTION
Employee engagement has become an important factor in business planning and human resource management [1]. Engaged employees are reported to make powerful contributions to desirable organizational outcomes, such as improved financial and operational benefits [27, 28, 67, 68]. To the extent that employee engagement can be influenced by technology, or measured through technology, it becomes an organizationally significant topic for studies in CHI and CSCW.

However, the CHI community is only just beginning to study employee engagement. Several papers have addressed engagement in broader terms, such as narrative engagement [55], flow [78], games [40], and commitment to specific workplace initiatives [20, 21]. A smaller group of studies has considered engagement in work, such as studies of mood and attention at work [16, 46, 77], and direct measurement of employee engagement from social media [75].

In general, organizational models of engagement have emphasized the importance of executive strategies [89, 92], managerial influences [18, 43], and workplace conditions that may be partially determined by executives and managers [14, 75, 89]. But, what about the employees themselves?

Towers Watson proposed an additional component, namely an algorithmically-specified set of individual responsibilities of each employee to manage her/his own engagement [84]. A broader set of influences was offered by Harter et al., who reported that close friendships at work were associated with greater engagement [27]; see also [23, 26]. In this paper, we expand on that basic finding with what we believe to be the first statistical evidence of the importance of social network factors as contributing to employee engagement in a large corporation. Using two years of survey results and social network relations from approximately 200,000 employees in a large corporation, we show that (a) an employee’s network position (centrality) and (b) her neighbors, (through homophily) can predict her engagement. This research adds to second social-media factors in what has been primarily a survey-based measurement practice. To the extent that social factors contribute to employee engagement, this work moves a business indicator more firmly into the realm of HCI workplace studies.

The remainder of this paper is organized as follows. First, we describe the context of our study, in terms of workplace and organizational priorities. This site-specific context sets up the themes of the paper, which we pursue in a literature review. Next we describe our data sources, which include primarily an engagement survey taken by over 200,000 employees, and the friendship ties among those employees in an internal social networking service. We present our network analyses, and discuss their implications for the use of social media in organizations, and for theories of employee engagement.
EMPLOYEE ENGAGEMENT AT LARGECORP

We studied the social network aspects of employee engagement at a large international company, LargeCorp (pseudonym for anonymous review). LargeCorp provides software and consulting services, and has offices in over 100 countries. Much of the work is both international and virtualized; virtual teams are the norm, and can vary from a few people to tens of thousands of people. To support the virtual teams, LargeCorp has implemented a suite of internal social software services for its employees, including a social-networking service, as well as many opportunities for structured or semi-structured online collaborations in the form of blogs, wikis, communities, and discussion forums. Similar services have been described in [9, 32, 38, 39, 49, 60, 65, 66, 75, 85, 86, 95]. In the case of LargeCorp, all of these social media, including a robust organizational social network (OSN), are contained behind the company’s firewall.

Like many commercial and government organizations, LargeCorp measures the engagement of its employees through an annual survey. The purpose of the survey is to measure progress of employee engagement and how employees are practicing the mission and values of the company. All regular employees of the company are invited to participate. In 2013, the overall participation rate was and 62% and in 2014 it was 55%.

Recently, LargeCorp has expanded its efforts from measuring engagement to predicting engagement. Predictions were made not for individual employees, but rather for larger groups of employees in what LargeCorp calls an “employee segment,” in support of detecting problems that affect many employees and that could be addressed through workplace programs that did not stigmatize individual members of the organization. Part of the motivation for this statistical treatment was to preserve employee anonymity. For all statistical analyses, records were aligned or joined on the basis of an MD5-encrypted version of the employee’s ID, which could not be decrypted by analysts or by management.

In work done before this paper, LargeCorp had investigated predictive factors such as employee sentiment (measured through employee social media) and employee demographics. In this paper, we examine additional, social network indicators that could expand the set of CHI-related factors related to this increasingly important organizational concept.

BACKGROUND

Studies of employee engagement fit within the broader context of workplace studies. Workplace studies have a long history in CHI and related conferences, in critical-historical terms [70], in specific work roles and disciplines [5], and in studies that deliberately cross and compare workplaces and their issues [31, 61]. Many western workplaces are organized hierarchically, with a simple, de facto network structure that takes the form of a tree (e.g., an “organizational chart”) [24, 30], or more recently a matrix. However, most organizations may also be described in terms of social networks that may or may not map onto the organizational chart [15, 53, 54, 66, 90]. CHI-and-related studies have shown initial and emergent phenomena in longer-term social networks [15, 17, 50, 51, 65, 76, 87] and social networking applications [34, 94].

Conceptions of Employee Engagement

In this paper, we examine the relationship of workplace social network attributes of individual (anonymized) employees, with their employee engagement. To explore this domain, we begin with general descriptions of employee engagement. After that, we introduce four potential influences upon engagement. We briefly review survey-based methods of measuring engagement, which help to define our dependent variable. Finally, we go into greater detail regarding social network aspects.

Definitions

Employee engagement has been described as “a heightened emotional connection [leading to] greater discretionary effort” ([23]; see also [75]); to “invest a high level of physical, cognitive and emotional resources on… work tasks…” [19]; or to “go the extra mile” [47]. Vance lists ten definitions from as many sources [88]. In a synthesis of seven models, Govender summarizes engagement as involving “one’s thinking, attitudes, feeling, behaviour and involvement in the organization” [26].

Despite strong commercial interest in employee engagement, academic research on this topic has been described as “a vast construct” [58] that is characterized by “a dearth of peer reviewed academic literature” [79]. Multiple researchers say that there is no consensus definition [1, 26, 44, 47, 69, 74, 79]. In her thesis, Seppälä differentiated among several distinct concepts: personal engagement, job engagement, work engagement, and employee engagement [73]. All of these concepts involve a positive attitude by employees to their work, and many research reports have tended to blend or combine them (e.g., [14, 18, 19, 23, 26, 35, 43, 44, 47, 58, 75, 79, 88, 89]). In commercial applications of the concept, the differences among these models tend to be treated as unimportant or imperceptible [75, 92, 93].
Several models propose an ordering or hierarchy of criteria for engagement. Robinson et al. describe engagement as being influenced by 11 factors, ranging from least important factors (e.g., job satisfaction and family-friendliness) to most important factors (e.g., performance appraisal, immediate management, and career opportunities) [64]. However, it seems likely that the ordering of those factors may vary by job role, as well as personal characteristics such as gender and national culture.

Several researchers have found consistency and convergence among these models. Schohat and Vigoda-Gadot differentiate employee engagement from somewhat related concepts, such as organizational commitment, job involvement, and organizational citizenship behavior [71]. Saks argues that employee engagement provides a unique blend of older concepts [68].

In summary, employee engagement has been associated with positive workplace outcomes [27, 28, 67, 68, 75]. Engagement is becoming a key internal organizational indicator [1]. Previous work has examined how social media text may be used to predict engagement [75]. We now explore predictive possibilities based on relationships revealed through a social networking service. Our exploration takes place within a more general model.

**Potential Influences on Employee Engagement**

If employee engagement is an important business attribute, then organizations are becoming interested to predict it [27, 28]. Figure 1 shows a summary view of four potential influences on the engagement of an employee.

**Context Engagement**

Traditionally, demographic characteristics combined with survey responses have been used to predict employee engagement [67,68]. Organizations have found that business circumstances cause certain country sites to have less engagement than others, or certain job roles to have less engagement than others. For example, employees in call centers are known to have higher turnover and less engagement [11]. While these attributes are not universal, and will vary by company and by industry, historical data and experience can allow an organization to build statistical models of employee engagement.

**Trait Engagement.** Much of the business practices regarding employee engagement are based on an implicit trait model. In this model, engagement is seen as a relatively stable attribute of a person, and is considered to be difficult to change – analogous to a personality trait, hence Trait Engagement. This perspective gained academic support in the work of Macey and Schneider [44], who analyzed engagement as deriving from an unmeasurable trait – which was manifested in an enduring attitude and would result in workplace behaviors.

**State Engagement.** If Trait Engagement is concerned with long-duration dispositions, then State Engagement is concerned with dynamic or rapidly-changing engagement perspectives. Using techniques from diary-based or experience-sampling studies, both Sonnentag et al. [80] and Breevaart et al. [8] showed that people’s degree of engagement may change markedly on a daily basis.

In a different approach, Shami et al. tested the ability to predict employee engagement (measured via a survey) based on text analysis of employees’ voluntary contributions to company-internal social media [75]. For an engagement survey conducted in one month, they showed that the social media texts from the previous month could predict the survey scores; however, prior months were not predictive. They proposed that employees’ state engagement might have a duration of a month or less.

**Social Engagement.** In this paper, we address a fourth possible influence. Most models of engagement involve primarily factors that can be influenced by executives or managers (e.g., [14, 18, 43, 75, 89, 92]). Towers Watson is an exception, suggesting that employees may carry some responsibility for their own engagement. We extend this line of thinking, to ask if the employee’s position in a company’s internal social network may also be predictive.
of her/his engagement. It the answer to this question is “yes,” then these network relations would become a fourth influence, Social Engagement, on an employee’s engagement.

Measuring Engagement through Surveys
First, however, we need a way to quantify each employee’s engagement. This is typically done through surveys.

Two influential academic models of engagement posit three factors each. Schaufeli and Bakker developed a survey called the Utrecht Work Engagement Scale (UWES), which measures three subscales based on concepts of vigor, dedication, and absorption [69]. Rich et al provided the Job Engagement Scale (JES), based on earlier work of Kahn [36], which measures three different subscales: physical, cognitive, and affective [63].

Commercially important models of engagement have also been introduced. Commercial work tends to be influenced in part by theory, and in part by the discovery of “actionable at [sic] (i.e., under the influence of) the work group’s supervisor or manager” [27], validated at a scale that is usually not possible for more academically-oriented research – i.e., hundreds of thousands of respondents across multiple companies. The Gallup “Q12” survey consists of twelve questions, which address themes of career growth and opportunity, recognition, resources for work, social relations at work, and alignment of the individual employee’s work with organizational goals [28]. The engagement portion of the Kenexa High Performance Engagement Model (HPEM) poses four questions, regarding pride in the company, satisfaction at work, advocacy/recruitment of others to join the company, and job retention [93].

There has been a tendency to decrease the size of surveys. The Utrecht survey began with 17 items, and then decreased to 15 items, and finally to 9 items [74]. Seppälä et al. showed that the 9-item scale had roughly the same factor structure and construct validity as its longer predecessors [74]. Commercial surveys tend to be shorter, in part because of the aggregate per-question cost in work-time when thousands of employees take a survey. The Gallup Q12 Engagement survey has twelve items [28]. A similar instrument by Kenexa contains eleven items, of our which only four items directly assess the topic of engagement [93]. LargeCorp took this simplification process a step further, and reduced the Kenexa survey to three items, which asked employees to rate their pride in working at LargeCorp, their satisfaction with LargeCorp as an employer, and the likelihood that they would recommend LargeCorp as an employer to a friend or relative.

Organizations and Social Networks
Social networks within organizations have been a focal point in organizational behavior studies [7] in the past thirty years. There are two primary categories of organizational social networks – formal and informal [2]. Formal networks generally encompass the strict, usually vertical organizational hierarchy whereas informal networks are those where the employees self-select their connections. The OSN that we investigate in this paper is an example of an informal network facilitated by the organization.

Newman and Park [57] argue that such types of self-selected social networks are different and unique in contrast to formal, imposed or inferred networks, because they contain two properties not found in other types of networks – network transitivity (or non-trivial clustering) and assortative mixing (or homophily).

These differences suggest the question - why are social networks an important structure to be studied within the context of the organization and why can they be useful in understanding and estimating employee engagement? Borgatti and Foster [7] in an early paper cite four areas of investigation in organizational social networks – social capital; expertise and information seeking; information diffusion or contagion; and environmental shaping (homophilic or social selection tendencies). Most areas of research have concentrated on one or more of these areas.

Nahapiet and Ghoshal provide further motivation for workplace studies of social networks [52]. They argue that employees may seek to maximize social capital within an organization because of three reasons. First, encouraging social capital often leads to the creation of intellectual capital. Second, organizations as institutions by definition are susceptible to the development of high levels of social capital. Third, organizations with more dense levels of social capital often have an advantage within their own markets in developing and disseminating intellectual capital.

Yuan et al. [96] and Fulk and Yuan [22] found that there are numerous ways in which employees satisfy their expertise or information seeking needs by leveraging their personal social networks within the organizations. Often, this depends on availability, accessibility, tie strength and the information system used. In fact, they [22] claim that in many contemporary organizations, enterprise social networking systems are the optimum platform for meeting information-seeking needs.

Strang and Soule [81] offer a comprehensive review of how information spreads through an organization. Among others, they mention two relevant points. First, network structure plays an important role in information diffusion. Thus, the way in which an employee chooses to position herself within the organization is an important predictor of information flow, and of the employee’s ability to receive and make use of information. Second, self-selected connections to other nodes are also a key predictor of diffusion.

In a similar vein, McPherson, Smith-Lovin and Cook [45] observe that individuals within organizations tend to self-
select their social ties to maximize social capital and to smooth information flow. Often, these ties are more in-group (homophilic) than out-group (heterophilic) because of pre-existing relationships, or the introduction to the group through an existing member. These type of powerful social forces tend to be homophilic in nature.

Summarizing this section, it seems clear that organizational social networks are important social structures because they contribute to many different phenomena, such as social capital, expertise seeking, and information diffusion. Moreover, unique features of such social networks are that groups of nodes cluster together because similarities between nodes usually nudge them to connect with each other (homophily). Employee engagement has been theorized to have social aspects [23, 26, 27, 84], but the implications of this phenomena have not been thoroughly studied.

We therefore ask:

RQ1: Is employee engagement predicted by node position in organizational social networks?

RQ2: Is employee engagement predicted by homophily in organizational social networks?

In the following section, we detail our methodological approach, the data collection process, a description of our data and the report relevant results.

**METHOD**

**Data Sources**

We combined three large workplace datasets:

- A social network of friendship relationships, recorded from LargeCorp’s social networking service. We discuss this network in the formal language of social network analysis, where people are “nodes” and friend relationships are “edges” that connect nodes. To protect employee privacy, all identities were anonymized through non-reversible MD5 transformation.

- Demographic data for each employee, provided by LargeCorp’s human resources organization, and anonymized in the same way as above. These data included locations, organizations, and human resources historical data such as time-in-title and performance ratings.

- Survey data from two years of employee engagement surveys, provided by the human resources organization, and anonymized in the same way as above. As described above, the survey used three core questions that address pride, satisfaction, and advocacy (likelihood to recommend LargeCorp as a good place to work).

We used the MD5-anonymized employee IDs to align these records. Our resulting dataset contained records for 209,471 employees for 2013 and 219,138 employees for 2014.

**Preprocessing**

We performed two pre-processing steps. First, we calculated several network centrality metrics from the edge-list data to determine the degree of importance of each node (employee) in the OSN. There are many network centrality metrics; we chose three that we felt best encapsulated the connectedness and importance of a node in our OSN, namely degree centrality, eigenvector centrality, and the clustering co-efficient. These metrics are formally defined in the next sub-section. Bonacich [6] recommended the inclusion of eigenvector centrality to measure the relative influence of the position of a node for large, complex networks while degree centrality is a simple measure of the number of friends of a particular node. Similarly, we chose the clustering co-efficient since prior work [56, 57] established that social networks exhibit a phenomenon of clustering based on preferential attachment of nodes. We note that feature selection of the optimum set of network centrality metrics for any given network is a separate, future research project and outside the scope of this particular report.

Second, consistent with current practices [93], we converted the mean score of the three survey items to a binary variable, where mean scores in the range 4-5 (inclusive) were coded as 1 (engaged) and mean scores below 4.0 were coded as 0 (unengaged). This was our dependent variable.

In the next three sub-sections, we first describe the different variables and predictors. Next, we discuss the statistical modeling approaches used to calculate both individual, as
well as network level effects on employee engagement. Finally, we detail the results and our inferences from them.

Variables and Measurements
We developed our model based on the following set of predictors:

Demographic and other Employee-related Variables
The descriptions and levels of measurement of demographic and other employee-related variables are as follows:

Salary Group: LargeCorp has an internal grouping system for employees in different salary groups. This consists of 15 groups. This is an ordinal variable.

Performance Review: LargeCorp also has an internal system of rating employees. This consists of 5 levels and is also an ordinal variable.

Time since Last Promotion: This represents the time (in months) since the employee was last promoted at work.

Tenure in the Company: This represents the time (in months) that the employee has been employed at the LargeCorp.

Age: This represents the biological age (in years) of an employee.

Gender: This is a binary variable (male=0/female=1) of an employee’s gender. LargeCorp did not provide alternative options to express gender.

Network Centrality Metrics
RQ1 predicted effects of node position in the social network. For broad coverage, we examined the following three network centrality metrics:

Degree centrality: This is a normalized (range = (0, 1)) measure of the number of nodes that a particular node is connected to in the OSN. [72]

Eigenvector centrality: This is a normalized (range = (0, 1)) measure of the relative influence of a node, given its position and connectivity in the OSN. [6]

Clustering Co-efficient: This is a normalized (range = (0, 1)) measure of how well a node tends to cluster together with immediate neighbors. [91]

Analysis
Our analysis has two parts. First, we used binary logistic regression to estimate the effect on engagement of an individual given her node position and demographical attributes. These results are presented in Table 2. Second, we estimated a social network-based effect by performing a standard homophily-influence analysis [33]. This analysis has two primary assumptions. First, it assumes that similar node connections influence each other greater than dissimilar ones. Second, it also assumes that the more commonalities two nodes have with each other, the more influence they have on each other. Therefore, for each node $p$ we compute the probability of a net homophily effect for engagement from all her neighbors as follows:

$$P(EE) = P(I) \times P(N) - P(I).P(N)$$

where

$u$ = total number of engaged neighbors
$v$ = total number of disengaged neighbors
$c$ = total number of common demographics with $p$
$t$ = total number of demographic groups
$n$ = $u + v$ (total number of neighbors)

Thus, the overall prediction of engagement for a given individual (I) as well as network (N) effects is computed by:

### Table 2. Results of the Logistic Regression Model for each of 2013 and 2014.

<table>
<thead>
<tr>
<th>Variables</th>
<th>2013</th>
<th></th>
<th>2014</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Std. β</td>
<td>OR</td>
<td>SE</td>
<td>Std. β</td>
</tr>
<tr>
<td>Salary Group</td>
<td>0.11</td>
<td>0.8</td>
<td>0.023</td>
<td>0.06</td>
</tr>
<tr>
<td>Performance Review</td>
<td>0.09</td>
<td>1.107</td>
<td>0.016</td>
<td>-0.12</td>
</tr>
<tr>
<td>Time since Last Promotion</td>
<td>-0.06*</td>
<td>0.859</td>
<td>0.018</td>
<td>-0.09*</td>
</tr>
<tr>
<td>Tenure in the Company</td>
<td>-0.14*</td>
<td>0.963</td>
<td>0.024</td>
<td>-0.12*</td>
</tr>
<tr>
<td>Age</td>
<td>-0.12*</td>
<td>0.717</td>
<td>0.033</td>
<td>-0.13*</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.03</td>
<td>0.952</td>
<td>0.014</td>
<td>-0.04</td>
</tr>
<tr>
<td>Degree</td>
<td>0.11**</td>
<td>1.118</td>
<td>0.001</td>
<td>0.07</td>
</tr>
<tr>
<td>Eigenvector</td>
<td>0.26***</td>
<td>1.614</td>
<td>0.016</td>
<td>0.28***</td>
</tr>
<tr>
<td>Clustering Co-efficient</td>
<td>0.09</td>
<td>1.031</td>
<td>0.021</td>
<td>0.19**</td>
</tr>
</tbody>
</table>

Model Fit Statistics

<table>
<thead>
<tr>
<th></th>
<th>2013</th>
<th></th>
<th>2014</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>3142</td>
<td></td>
<td>3318</td>
<td></td>
</tr>
<tr>
<td>BIC</td>
<td>2692</td>
<td></td>
<td>2958</td>
<td></td>
</tr>
</tbody>
</table>

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$
To assess these predictions, we compared correct predictions vs. incorrect predictions via confusion matrices. The confusion matrix for the computation of $P(I)$ as well as $P(EE)$ for each year of analysis (2013 and 2014) after 10 fold cross validation is presented in Table 3.

**RESULTS**

Table 1 lists summary statistics describing the major properties of the OSN across both years (2013, 2014). We can get a holistic idea of the structure and size of the OSN across both years. There are two things to note here that we will allude to further in this section. First, comparing 2013 and 2014, we observed a net increase of 415,319 edges (friend relationships) and 9667 nodes (employees) in the OSN. Second, LargeCorp reorganized itself internally during this period which was likely to have impact on its OSN structure.

**Individual Level Effects (RQ1)**

We used binary logistic regression, predicting engagement on the basis of demographics and OSN centrality measures, to estimate the individual level effect. Our dependent variable was the binary outcome obtained from the engagement survey (engaged vs. unengaged), and our predictors were the demographic and social network metrics described above in the “Variables and Measurements” section. The results of our regression analysis are shown in Table 2.

<table>
<thead>
<tr>
<th>Predicted</th>
<th>2013</th>
<th></th>
<th>Predicted</th>
<th>2014</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0 38,819 (44.9%)</td>
<td>36,421 (31.4%)</td>
<td></td>
<td>0 35,414 (46.2%)</td>
<td>39,716 (27.9%)</td>
<td></td>
</tr>
<tr>
<td>1 47,637 (45.1%)</td>
<td>79,569 (69.6%)</td>
<td></td>
<td>1 41,239 (43.8%)</td>
<td>102,637 (72.1%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Confusion Matrices before and after Network Effects

To assess these predictions, we compared correct predictions vs. incorrect predictions via confusion matrices. The confusion matrix for the computation of $P(I)$ as well as $P(EE)$ for each year of analysis (2013 and 2014) after 10 fold cross validation is presented in Table 3.

**RESULTS**

Table 1 lists summary statistics describing the major properties of the OSN across both years (2013, 2014). We can get a holistic idea of the structure and size of the OSN across both years. There are two things to note here that we will allude to further in this section. First, comparing 2013 and 2014, we observed a net increase of 415,319 edges (friend relationships) and 9667 nodes (employees) in the OSN. Second, LargeCorp reorganized itself internally during this period which was likely to have impact on its OSN structure.

**Individual Level Effects (RQ1)**

We used binary logistic regression, predicting engagement on the basis of demographics and OSN centrality measures, to estimate the individual level effect. Our dependent variable was the binary outcome obtained from the engagement survey (engaged vs. unengaged), and our predictors were the demographic and social network metrics described above in the “Variables and Measurements” section. The results of our regression analysis are shown in Table 2.

Broadly, our results are consistent across both years of analysis. Demographics, in both years, explain roughly 50% of the overall variance. In 2013 and 2014, among demographic variables, “Time since Last Promotion” (OR$_{2013}$=0.859, OR$_{2014}$=0.556), “Tenure in the Company” (OR$_{2013}$=0.963, OR$_{2014}$=0.896) and “Age” (OR$_{2013}$=0.717, OR$_{2014}$=0.834) were significant predictors of an increased likelihood of engagement. A close inspection of all the demographic variables starts to reveal a picture of an ideal engaged employee. This hypothetical employee has been with the organization for a short period of time (1-3 years), earns a mid-level salary, was promoted less than a year ago, and usually receives a high performance review. Altogether, this suggests that engagement at LargeCorp is driven only partly (~50% of explained variance) by demographic and other workplace related factors. Our question then becomes: can we improve on this prediction by adding network centrality measures to our predictors?

**RQ1** predicted that node position in the network would be an important predictor of employee engagement. In 2013, degree centrality and eigenvector centrality were significant predictors in this model. This result points towards an intuition that employees arrange themselves in specific ways in this OSN to achieve their enterprise objectives. Particularly, a modest number of friends, and the relative influence of their position in the network, are important predictors of engagement. This situation slightly changes in 2014, where eigenvector and clustering co-efficient are the significant network centrality predictors. This suggests two outcomes of interest: First, relative node position important in the OSN; Second, the tightness of the nodes’ connections with each other is equally important.

We found a difference in predictive factors from one year to the next (i.e., degree centrality was a significant predictor for 2013 only, and clustering coefficient was a significant predictor for 2014 only). One possible explanation may be found in Table 1, which shows that there is a difference between the OSN in 2013 and 2014. Comparing between 2013 and 2014, the net edge-to-node ratio is 42.96 which suggests that a disproportionate number of edges have been added in 1 year as compared to the number of nodes. Since, the average non-normalized degree centrality is 9.91 in 2013 and 11.34 in 2014, we suspect that this was related to LargeCorp’s internal restructuring which took place at the end of 2013, which may have resulted in greater cohesion between people. Moreover, the social networking service is relatively new (deployed in 2009) and employees at LargeCorp are encouraged to use the features by
management. This could also account for employees using the platform more across years.

Summarizing this section, we find that individual level effects (consisting of demographics and network centrality measures) both matter significantly in estimating employee engagement. This answers RQ1.

**Network Level Effects (RQ2)**

The confusion matrices of Table 3 give us an idea of the increase in predictive power and explained variance of the model after taking into account the influence that the neighbors of a node exerts upon her.

In the 2013 OSN, we find that the addition of homophilic influences to individual level effects increase predictive power of true positives by 8.3% and true negatives by 10.5%. Similarly, in the 2014 OSN, we find the predictive power of true positives increased by 7.7% and true negatives by 8.5%. These are important points to note for two reasons. First, these increases are moderately high and second, they are in line with effect sizes in existing work [45] arising from homophily. This gives us a relatively high degree of confidence that homophily (or social selection) is an important predictor in estimating EE in LargeCorp. In other words, how you choose your friends in your OSN matters significantly towards your EE.

**RQ2** asked if homophily in OSNs have some influence on EE. Our homophily analysis suggests that indeed, similarity among neighbors in OSNs affect EE significantly. This answers RQ2. The existence and increase in homophily, in significant amounts, could also be one explanation of the statistical significance [57] of the clustering co-efficient metric that we found in Table 2. We believe that, in the context of estimating EE, this is a novel and important finding in the organizational network literature which we discuss in greater detail in the subsequent section.

**DISCUSSION**

Our paper has several implications that we group into three broad themes. First, we focus on the implications for social networks, particularly organizational social networks. Second, we describe implications for research on employee engagement and for workplace studies. Finally, we also comment on the broader implications of our work with existing research situated in the broader HCI community.

**Implications for Organizations and Social Networks**

Prior work [7] has found that people in organizations seek to forge relationships with other employees primarily because of four factors – social capital management; expertise and information seeking responsibilities; disseminating information through the network in an optimum way and for seeking to connect with other, similar employees to create better work environments. We found that previous models of employee engagement (e.g., [14, 18, 43, 59, 64, 89, 92]) do not take into account social influences upon employees within the organization.

Moreover, employee engagement has been theorized [19, 47] to be an important contributor to job satisfaction and workplace pride, which themselves are important factors [25,42] for the continued growth of the organization.

Our findings are situated at the confluence of these areas of prior work. To summarize, we found that employee engagement is dependent on network position and centrality and that, connections to homophilous friends influence employee engagement more than connections to heterophilic friends. Thus, it does not seem too farfetched to propose that social factors do indeed influence EE within organizations.

Employees at LargeCorp are offered access to an intra-organizational social network and significantly encouraged to use this OSN for most, if not all, work activities. This platform is similar (in its core features) to existing platforms in the public Internet i.e. Facebook, Google+ etc. Presumably, employees at a technological focused company have had some prior experience with social networking applications. This is then, a natural way for them to focus on their organizational social objectives [7] in line with Fulk and Yuan [22] who pointed out that in the modern organization, enterprise social networking platforms offer employees one of the best ways to achieve their organizational objectives.

Homophily in organizational social networks is an ever-present and undeniable phenomenon [45, 50, 62]. Most studies of homophily tend to focus on demographic variables such as race, gender etc. (e.g., [45, 62]). However, in this paper we focused mostly on homophily variables that are characteristics of a career or job role – independent of true demographics such as tenure in the organization, salary, performance review. These homophilic factors are shown to have a stronger influence on EE. Thus, at least for estimating EE, it is actually similarity about the organization structure and career or job role that is more important than the absolute demographics of an employee. This has not been studied in depth in previous studies of organizational social networks.

We can examine and infer from our findings from yet another perspective. We have spent the last few paragraphs discussing employee engagement. However, we should also address the other side of this coin – employee disengagement. Prior work suggests [1] that employee disengagement is also growing problem in large organizations. RQ1 and RQ2 suggests that social factors affect employee engagement but they have an equal effect on employee disengagement as well. We have discussed earlier that surrounding oneself with similar, engaged employees will generally have a positive effect on an employee’s own engagement.
However, being surrounded by similar, disengaged employees will have a negative effect on EE. This is an important factor of note in organizational social networks. While OSNs are a powerful way to achieve organizational objectives [22], they can also serve as an equally, optimum way to foster employee disengagement. We established earlier that similar people tend to cluster together in social networks [57] and our results, particularly in reference to Tables 2 and 3 support this hypothesis. While one obvious design implication from this work is to foster communities of engaged employees in OSNs, another equally important implication is the detection of disengaged communities. Managers and practitioners could then promote workplace policies designed to support the injection of more engaged employees within these identified communities or to improve these stagnant communities by pinpointing and alleviating core areas of concern that are the root causes of employee disengagement. This might best be achieved through a qualitative inquiry within the organization.

**Implications for Social Computing/HCI**

As more organizations explore the use of social media in the workplace [3, 10, 98], we have growing opportunities to study online social ties in relation to people’s work activities and relationships. Our results show that social network position may have important implications for employees’ experiences with what might be called the “organizational interface” to their workplaces. Increasingly, that organizational interface takes the form of corporate intranet portals and other forms of mediated communications. This paper suggests experiments in making the organizational interface more social in nature, through tools for social exchange and social networking.

Organizations typically use surveys for measuring employee engagement [41]. We know that survey response rates vary. A review of 69 email surveys finds that the response rate averaged around 40% [13]. While certain organizations may have higher response rates, there will be a population of employees from which the organization will not receive survey responses. Our results point towards the possibility that social network characteristics can be used to infer engagement levels of employees that do not respond to employee engagement surveys. Imputations of ‘missing values’ for these employees and subsequent validation could be a fruitful avenue of future research.

Organizations have typically entrusted first line managers with the engagement of their employees. Many of the employee engagement related change management actions focus on first line managers. Our results point towards other social influences that can influence engagement, namely network position, and social network relationships. Organizations can take advantage of this by focusing on enriching social connectedness. Social connectedness could be enhanced by deploying social networking tools that are easy to use and allow discovery of others across the organization with similar tastes, interests, job roles, projects etc. Recommender systems could be built to suggest names of other employees an employee should connect with [12]. Tools such as expertise locators [83] could allow employees to find others with similar expertise and foster communities of practice. Such social relationships can augment the efforts of first line managers to keep employee engagement up.

Our results argue in favor of deploying social network systems in the workplace. While gradual acceptance of social networking technologies is increasing in the workplace, there is still skepticism of the value of such technologies for work [4]. Our results extend prior research that shows the use of internal corporate social networking sites is positively associated with bonding relationships, a sense of corporate citizenship, interest in connecting globally, and access to people and expertise [82]. The relationship between social network connectedness and employee engagement is another argument in favor of using corporate social networks internally.

**LIMITATIONS AND FUTURE WORK**

Of course, our data were from a single company, which had a high degree of participation in a relatively rich set of internal social media resources. Similar results for other organizations may depend on the extent of their use of social networking tools.

We noted that the friendship-based social network of LargeCorp was only six years old, and may still be maturing. It will be interesting to see if survey results in future years show similar patterns to what we found for 2013 and 2014, or if the maturity of the social network leads to different phenomena.

Finally friendship-based social links are only one of several types of links that can be derived from LargeCorp’s social software environment. Other possible link-types include: co-membership in online communities, and various reply-to links that can be easily derived from discussion forums (topic and reply) or blogs (post and comment), or even a sequence of revisions to a wiki page. Different modalities of links can be weighted together to estimate tie strength between each connected node. This will give us a more sophisticated and nuanced way to understand the effect of homophily on the EE of a particular node. We hope to compare the social network relationships that can be computed from these types of links, with the friendship-based effects that we reported in this paper.

**CONCLUSION**

In this paper we reported the results of an empirical study in a large, multinational company (LargeCorp) on approximately 200,000 employees. We examined the relationship between node centrality, neighbor influence (based on homophily) and demographic factors to estimate their influences on EE. This study provides several contributions. First, we show that node position and centrality (in OSNs) affects EE in organizations. Second,
we show that homophily in OSNs also affect EE. Third, we inform the current theoretical work on EE by addressing the social component of employee engagement. Third, we connect the impact of this work with HCI, workplace studies and social network analysis. Finally, while our analysis is limited to a single company, the generalizability of our inferences from a large scale study on 200,000 employees provides, in our opinion, sufficient external validity for our work to be valuable to other researchers and practitioners to build and extend upon.

REFERENCES


